

Horn Dániel

## Az iskolai hozzáadott érték mérése

### Bevezetés

Az elmúlt években, évtizedekben egyre nagyobb hangsúlyt kap az iskolák elszámoltathatósága. A 2001-es amerikai *No Child Left Behind* törvényt<sup>1</sup> már azzal a feltevéssel fogadták el, hogy iskolai, illetve egyéni standardok felállítása és ezek elérésének rendszeres mérése, illetve ezekhez kapcsolódó ösztönzők kialakítása növelni fogja a tanulói – és egyben tanári – teljesítményeket.<sup>2</sup>

A 2008-as évtől kezdve az Országos kompetenciamérés is kihasználja azt a lehetőséget, hogy minden közoktatásban tanuló diák rendelkezik oktatási azonosítóval. Ez a kódszám, illetve az ebből számított mérési azonosító lehetőséget teremt a tanulók kompetenciamérés-eredményeinek évek közötti egyéni szintű összekötésére, ami, többek között, lehetőséget teremt a hozzáadott érték típusú eredményességi mutatók számítására. Jelenleg az iskolák főleg keresztmetszeti szintű adatokon becsült „iskolahatásokat”<sup>3</sup> kapnak kézhez. Ez különösen azért probléma, mert ezen iskolahatások nem veszik figyelembe a diákok előző teljesítményét, vagyis jobbnak

<sup>1</sup> Elérhető: <http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf> [Letöltés dátuma: 2015.04.30.]

<sup>2</sup> Feltehetőleg hasonló premisszán alapul az EMMI 2014/35-ös rendelete, illetve különösen ennek 9§-nak 10. bekezdése is, amely az Országos kompetenciamérések alapján készült jelentések figyelembevételével rendeli el iskolai intézkedési tervek kialakítását és ezek végrehajtását.

<sup>3</sup> Iskolahatás alatt az iskola valamilyen statisztikai módszerrel számított eredményességét értjük a tanulók valamilyen jellemzőire (pl. teszteredményeire). Vagyis a továbbiakban az iskolahatás az összes olyan mutatószám összefoglaló neve, beleértve a hozzáadott-érték típusú mutatókat is, amely az iskola tanulóira gyakorolt hatását próbálja számszerűsíteni.

*A tanulmány részben támaszkodik az MTA-KTI A közoktatás teljesítményének mérése-értékelése, az iskolák elszámoltathatósága 2008–2011 között futó programjának ACC 1503. számú produktumára.*

mutatják a képességek szerint szelektáló iskolákat.<sup>4</sup> A telephelyi jelentések ugyan tartalmazznak egy „komplex modellt”, amely egy bizonyos típusú hozzáadott érték mérés eredményeit mutatja, ám ez csupán egy indikátor a több tucat között. Hozzáadott érték típusú indikátorból azonban nem csak egyféle létezik.

Jelen tanulmánynak nem célja e külső mérések és az ehhez kapcsolt ösztönzők jogosságának vagy eredményességének elemzése. A tanulmány bemutatja az elméletileg lehetséges hozzáadott érték típusú indikátorok alapverzióit és néhány nemzetközi példát is hoz ezekre. Továbbá kísérletet tesz arra, hogy magyar adatokon megvizsgálja, valóban más eredményeket kapnának-e az iskolák, ha nem keresztmetszeti, hanem több időpontban mért, tanulói szinten összekötött adatokon becsült, hozzáadott érték típusú eredményességi mutatókat kapnának kézhez. Ez a kutatás csupán egy *pilot*, amely nem alkalmas iskolai szintű következtetések levonására, de az oktatáspolitikai döntéshozóknak kulcsfontosságú következtetésekkel szolgálhat.

A tanulmány első fele az elméleti, míg a második az empirikus fejezeteket tartalmazza. Az elméleti rész érvel a hozzáadott érték típusú mutatók fontossága mellett, és kifejti néhány alaptípusát a hozzáadott érték típusú modelleknek. A hozzáadott érték típusú modellekből négy, a gyakorlatban is felhasznált modell is bemutatásra kerül. Az empirikus rész a jelenleg használt Országos kompetenciamérést mutatja, továbbá az elméletben bemutatott eredményességi mutatókat veti össze az iskolajelentések eredményeivel.

### Hozzáadott érték-modellek az elméletben

Hogy mit tekintünk az iskola teljesítményének, egyaránt definíciós és módszertani kérdés (lásd még e kötetben *Szemerszki Marianna* írását). Az elszámoltathatósági rendszerek leginkább a tanulói teszteredményeket használják fel kimenetként. Természetesen az e tesztek által mért tanulói teljesítmények, illetve az ezekből számított iskolai teljesítmény csupán kiragadott mutatói az iskola valódi teljesítményének, ám feltehetőleg a rendelkezésre álló mérések közül még mindig a legpontosabbak, legmegbízhatóbbak. Ezen teszteredmények nagy előnye, hogy általában fontos alapképességeket mérnek (például szövegértés, matematika, természettudományok), ráadásul folytonos mutatók (szemben például a diszkrét to-

<sup>4</sup> A hozzáadott érték típusú modellek előnyeiről és hátrányairól lásd például *Ladd és Walsh* (2002), vagy *Kane és Staiger* (2002), a felhasznált termelési függvényről pedig *Todd és Wolpin* (2003), vagy *Dolton* (2002) tanulmányait.

vább tanul/nem tanul tovább, vagy a kategorikus dolgozik/munkanélküli/inaktív mutatókkal), ami módszertanilag megkönnyíti az indikátorok számítását (lásd még e kötetben *Nahalka István* írását).

A tanulmány e teljesítménymutatókból számított iskolai eredményesség indikátorainak sokféleségét hangsúlyozza, és nem foglalkozik a különböző eredményességfelfogásokból adódó eltérésekkel. Egy adott tanulói teljesítménymutatóból is nagyon sok eredményességi indikátort lehetséges előállítani. A gyakorlatban is szinte feltérképezhetetlen variációkban fordulnak elő egyszerű keresztmetszeti adatokon, vagy több időpontban történő méréseken nyugvó hozzáadott érték típusú eredményességi mutatók, amelyek ugyanazon kimenetet használnak fel. A tanulmány e mutatók leggyakoribb és az oktatáspolitikai döntéshozók számára talán leginkább felhasználható variánsait mutatja be.

A *keresztmetszeti adatokon nyugvó*, iskolai eredményességet leíró indikátoroknak két nagyon fontos előnyük van. Az *első*, hogy szinte azonnal, a felmérést követően és általában sokkal nagyobb arányban rendelkezésre állnak, mellyel szemben a hozzáadott érték típusú mutatókhoz minimum két időpontban szükséges megfigyeléseket végezni egy adott iskolában, és így az ezeken alapuló visszajelzés is sokkal lassabb. Ráadásul a két időpontban végzett felmérések elkerülhetetlenül nagyobb válaszmegtagadással vagy mintakopással járnak, ami szintén gyengítheti a mérés megbízhatóságát. A *másik előnye* e mutatóknak az egyszerű értelmezhetőségük. A keresztmetszeti adatokon nyugvó nyers mutatók általában egyszerű iskolai átlagok, amik – egy jellemző példával szemléltetve – például azt mutatják meg, hogy egy adott iskolában a tanulók hány százaléka teljesít egy megadott minimumszint alatt. E mutatóknál nincs szükség bonyolult statisztikai vagy ökonometriaival számításokra, s így az indikátorok visszacsatolási funkcióját is feltehetőleg könnyebben betöltik. Ráadásul adott indikátorok közpolitikailag is fontos üzenetet hordozhatnak: az előbbi példa esetében például azt, hogy függetlenül az iskola társadalmi összetételétől, programtípusától, fenntartójától vagy egyéb jellemzőitől minden tanulótól (illetve a tanulók megadott arányától) elvárt, hogy a minimumszint felett teljesítsen. Nyilvánvalóan némely iskolának egy ilyen indikátornak való megfelelés nem jelent kihívást, míg egy másiknak igen. Ugyanakkor éppen ez a keresztmetszeti adatokon nyugvó indikátorok egyik legnagyobb hátránya is: a tanulók jellemzői, illetve az iskolától független egyéb jellemzők nagymértékben meghatározzák, hogy egy adott iskola miként teljesít. Egy átlagosan jobb képességű tanulókat oktató iskolában várhatóan akkor is sokkal nagyobb arányban fognak a tanulók a megadott teljesítménymínimum felett teljesíteni, ha az iskola semmivel nem jobb (a szó általános értelmében), mint egy másik iskola, ahol a tanulók kevésbé jó ké-

pességűek. Épp ennek a torzításnak a csökkentése miatt fontos a hozzáadott érték típusú mutatók számítása.

Ahhoz, hogy a hozzáadott érték típusú mutatók előnyét be lehessen mutatni, be kell vezetni az oktatási témákkal foglalkozó közgazdászok (oktatásgazdászok) által használt oktatási termelési függvény fogalmát.<sup>5</sup> A termelési függvény azt feltételezi, hogy az iskola egy meghatározott kimenetét (például tanulói tesztpontszámok) különféle, az iskolától független, illetve az iskolától függő változók befolyásolják. *Todd és Wolpin* (2003) részletesen összefoglalja a különböző termelési függvények becslésének módszereit, feltevéseiket és minimális adatigényüket. Cikkük alapján *Dolton* (2002) a következő egyszerűsített modellt mutatja be.

Feltéve, hogy a tanuló iskolai teljesítménye az egyéni veleszületett képességek és tulajdonságok, az iskolai és a családi erőforrások függvénye, a következő általános termelési függvényt írhatjuk fel:

- (1)  $Y_{ijkt} = g(X_{it}, \Sigma S_{ijkt}, \Sigma F_{it}, \mu_i)$ , ahol
- $Y_{ijkt}$  a  $t$ -edik évben,  $j$ -edik iskolába,  $k$ -adik osztályba járó,  $i$ -edik tanuló teljesítménye;
  - $X_{it}$  a  $t$ -edik évben az  $i$ -edik tanuló minden olyan személyes jellemzője, amely befolyásolhatja a teszteredményét;
  - $S_{ijkt}$  a  $t$ -edik évben,  $j$ -edik iskolába,  $k$ -adik osztályba járó,  $i$ -edik tanuló rendelkezésére álló iskolai erőforrások;
  - $F_{it}$  a  $t$ -edik évben az  $i$ -edik tanuló minden családi jellemzője, erőforrásai;
  - $\mu_i$  az  $i$ -edik tanuló veleszületett képességei.

Azaz az adott évben megfigyelt tanulói jellemzők, az adott és az azt megelőző években megfigyelt iskolai erőforrások és családi jellemzők, erőforrások, illetve a tanuló veleszületett képességei mind befolyásolják valamilyen módon az adott időpontban megfigyelt tanulói teljesítményt.

Így az első időszakra például a következő egyszerű függvényt becsülhetjük (feltéve, hogy az összefüggés a tanulói teljesítmény és az inputok között lineáris és additív):

$$(2) \quad A_{ijk1} = \alpha_1 X_{i1} + \beta_1 \Sigma S_{ijk1} + \delta_1 \Sigma F_{i1} + \gamma_1 \mu_i + \varepsilon_{i1} + u_{j1},$$

<sup>5</sup> *Horn* (2011) felhasználásával.

ahol  $\alpha, \beta, \gamma, \delta$  becslési paraméterek, míg  $\varepsilon_i$  és  $u_j$  egyéni és iskolai szintű nem megfigyelt változók hatását mutatja.

Az ilyen termelési függvényekkel végzett elszámoltathatósági vizsgálat célja általában az  $u_j$ , azaz az iskolai szinten nem megfigyelt változók hatásának becslése. Ez a hatás egy úgynevezett reziduum-, vagyis maradék-hatás. Az  $u_j$  minden olyan iskolaszintű hatást magában foglal, amit az egyéb, a modellbe bevont változók nem magyaráznak. Ez a szám mutathatja meg, hogy egy adott időszakban hogyan teljesített az iskola az átlaghoz képest, figyelembe véve az iskolába járó tanulók képességeit és tulajdonságait, illetve az iskola és a család erőforrásait. A becslés igencsak adatigényes, hiszen például ismertnek feltételezi a diák veleszületett képességeit, illetve az elmúlt időszakok iskolai és családi ráfordításait, amelynek mérése nem egyszerű feladat.

A közgazdászok e nagy adatmennyiség kikerülésére a következő módon becsült, úgynevezett „hozzáadott-érték” vagy „student-gain” (tanulói-nyereség) típusú becslést szokták elvégezni. Elméletileg hasonlóan az első időszakhoz, a második időszakra is meg lehet becsülni a termelési függvényt:

$$(3) \quad A_{ijk2} = \alpha_2 X_{i2} + \beta_2 \Sigma S_{ijk2} + \delta_2 \Sigma F_{i2} + \gamma_2 \mu_i + \varepsilon_{i2} + u_{j2}.$$

A kettő különbségét véve ezt kapjuk:

$$(4) \quad \Delta A = \alpha_2 X_{i2} - \alpha_1 X_{i1} + \beta_2 \Sigma S_{ijk2} - \beta_1 \Sigma S_{ijk1} + \delta_2 \Sigma F_{i2} - \delta_1 \Sigma F_{i1} + \gamma_2 \mu_i - \gamma_1 \mu_i + \varepsilon_{i2} - \varepsilon_{i1} + u_{j2} - u_{j1}.$$

Ami az alábbi feltevésekkel a következő formára egyszerűsíthető:

$$(5) \quad \Delta A = \alpha X_i + \beta S_{ijk} + \delta F_i + \varepsilon_i + u_j.$$

Feltételezések:

1. A tanulók személyes jellemzői nem változnak az idő múlásával ( $X_{i1} = X_{i2}$  így  $\alpha_2 X_{i2} - \alpha_1 X_{i1} = \alpha X_i$ ). Ez természetesen azt jelenti, hogy olyan jellemzők, mint a tanulói motiváció vagy erőfeszítés sem változik. Ez nyilvánvalóan nem felel meg a valóságnak, így eléggé erős feltételezés.
2. Az iskolai erőforrásokat egy adott időpontban megfelelően meg lehet figyelni, így  $\beta_2 \Sigma S_{ijk2} - \beta_1 \Sigma S_{ijk1} = \beta S_{ijk}$ , vagyis az előző ráfordítások hatása kiesik az egyenletből.

3. A családi ráfordításokat egy adott időpontban megfelelően meg lehet figyelni, így  $\delta_2 \Sigma F_{i2} - \delta_1 \Sigma F_{i1} = \delta F_i$ , vagyis az előző ráfordítások hatása kiesik az egyenletből.
4. A tanulói veleszületett képességek hatása az idővel nem változik, azaz  $\gamma_2 = \gamma_1$ , így a veleszületett képességek hatása kiesik az egyenletből. Ez a feltevés is erős, hiszen valószínű, hogy az életkor növekedésével a gyermek képességeinek hatása változik.
5. A család és az iskola általi ráfordítások nem függnek össze a tanuló(k) korábbi teljesítményével, azaz  $S_2$  és  $F_2$  nem korrelál  $Y_0$ -al illetve  $\mu_i$ -vel. Tekintve, hogy például várhatóan a szülők sokkal több energiát fordítanak gyermekük oktatására mint korábban, ha kiderül, hogy rosszul teljesít az iskolában, ez a feltevés sem felel meg teljes mértékben a valóságnak.

Bár a felvételek igencsak erősek, ha elfogadjuk ezeket, az iskolai hatás ( $u_i$ ) becsléséhez már reális mennyiségű adat is elegendő. Ez esetben szükség van legalább két időpontban mért, tanulói szinten összekötött adatokra a tanuló iskolai teljesítményéről, illetve legalább egy megfigyelésre az iskola és a család ráfordításairól, és a tanuló tulajdonságairól. Ráadásul, ha az utóbbi változókat mindkét időpontban megfigyeljük, azaz részletesebb adatok állnak rendelkezésre, a feltételezések egy része feloldható.

### Hozzáadott érték-modellek a gyakorlatban

A hozzáadott érték típusú modellek bemutatását a legegyszerűbb lineáris modellekkel kezdeni. Ezek mind felvételeikben, mind adatigényében viszonylag egyszerűek, és pont ezért könnyen átláthatóak és relatíve könnyen megérthetőek. A modellek néhány, gyakorlatban is alkalmazott, nem lineáris specifikációját a következő alfejezet tárgyalja.

A hozzáadott érték típusú modellek közös tulajdonsága, hogy minimum két időpontban megfigyelt, tanulói szinten, az időpontok között összekötött adatokat használnak fel. Egy jelentős leegyszerűsítéssel élve a modelleket két dimenzió alapján lehet csoportosítani:

1. az egyik dimenzió, hogy a múltbeli teljesítmény-pontszámokat függő vagy független változóként veszi-e figyelembe,
2. a másik pedig, hogy az iskola hatásának random vagy fix hatást becsül.

Az *első dimenzió* szerint 3 típust különböztethetünk meg: (1) a múltbeli teljesítményt, mint független változót felhasználó („*kiigazított*”) modellek, (2) a több

időpontban megfigyelt kimenetek különbségét felhasználó („*fejlődés-*”) modellek, illetve (3) ezek kombinációja („*kombinált*” modellek). A *másik dimenzió* szerint – hogy random vagy fix hatással becsülik-e az iskola hatását, vagyis a hozzáadott értéket – szintén *három eset* lehetséges: random hatás becslésénél lehet (1) az egyéni hibák iskolai átlagát vagy (2) az iskolai random hatást vizsgálni, míg fix hatás esetében (3) az iskolai fix hatások felelnek meg a hozzáadott értéknek.

Az 1. táblázat összefoglalja a két dimenziót, és az általuk meghatározott hozzáadott érték mutatókat, illetve azt a jelölést, amelyet a továbbiakban a tanulmány alkalmaz.

1. táblázat

A tanulmányban számolt indikátorok és ezek rövidítései

		Keresztmetszeti adatokon alapuló iskolahatás	Hozzáadott érték típusú mutatók		
			„kiigazított”	„fejlődés”	„kombinált”
<b>Random hatás</b>	Egyéni hibák iskolai átlaga	$IH_E$	$HE_{E1}$	$HE_{E2}$	$HE_{E3}$
	Iskolai random hatás	$IH_I$	$HE_{I1}$	$HE_{I2}$	$HE_{I3}$
<b>Fix hatás</b>	Iskolai fix hatás	$IH_F$	$HE_{F1}$	$HE_{F2}$	$HE_{F3}$

Az alábbiakban először az első dimenzió logikája kerül bemutatásra a legegyszerűbb, egyéni hibák iskolai átlagát figyelembe vevő modell segítségével, majd az iskolai random és fix hatás különbsége kerül bővebb kifejtésre.

### Random hatás-modellek – az egyéni hibák iskolai átlaga

A három felsorolt modell típus közül ez a legegyszerűbb. Az iskolahatás a tényleges és a modell által becsült érték közötti különbség (az egyéni hiba) iskolai átlaga; vagyis ugyanaz a logika, mint az Országos kompetenciamérés telephelyi jelentéseiben (lásd alább). A modell feltételezi, hogy az egyéni hiba normális eloszlású, egyéni szinten várható értéke 0 és szórása konstans, illetve független minden modellbe illesztett kontrollváltozótól.

**A) „Kiigazított” (covariate adjustment) modell**

A modell legáltalánosabb formája:

$$(6) \quad y_{ijt} = a_0 + a_1 y_{ij(t-1)} + bX + e_{ij},$$

ahol a fentebb bemutatott jelölések mellett  $t$  az idő jele. A hozzáadott érték ( $HE_{E1}$ ) a modell által becsült és a valós érték közötti eltérés, azaz

$$(7) \quad HE_{E1} = \text{átlag} (y_{ijt} - \hat{y}_{ijt}) = \hat{e}_{ij},$$

ahol  $\hat{y}_{ijt}$  az adott diákra becsült érték.

A modell tehát azt feltételezi, hogy a tanuló adott évi teljesítménye függ az előző teljesítményétől, és egyéb egyéni (családi) és iskolai jellemzőitől.<sup>6</sup> Ezek segítségével lineáris becsléssel közelíti a jelenbeli teljesítményét. Amennyiben egy iskolában a tanulók átlagosan ennél, azaz az elvártnál jobban teljesítenek, az adott iskola hozzáadott értéke pozitív, ellenkező esetben negatív.

**B) „Fejlődés” (gain) modell**

Ennél valamelyest restriktívebb feltételezésekkel él a következő modell:

$$(8) \quad (y_{ijt} - y_{ij(t-1)}) = a_0 + bX + e_{ij}.$$

A hozzáadott érték ( $HE_{E2}$ ) itt is a hibatag iskolai átlaga:

$$(9) \quad HE_{E2} = \hat{e}_{ij}.$$

A két modell közötti legfontosabb különbség, hogy míg  $HE_{E1}$  azt feltételezi, hogy a múltbeli teljesítmény és az egyéni, illetve iskolai jellemzők befolyásolják a jelenbeli teljesítmény szintjét, addig a  $HE_{E2}$  szerint az egyéni és iskolai jellemzők a teljesítménynövekedés ütemét befolyásolják. Azaz, míg a „kiigazított” modell megengedi, hogy az azonos egyéni vagy iskolai jellemzővel rendelkező diákok másképp fejlődjenek az adott periódus alatt, addig a „fejlődés”-modell azt feltételezi, hogy az azonos egyéni, illetve iskolai jellemzőkkel rendelkező diákok azonos ütemben fej-

<sup>6</sup> Az, hogy a gyakorlatban a modell milyen egyéni és iskolai változókat vesz/vehet figyelembe, külön tanulmányt érdemelne, és az adott kérdésben az oktatáspolitikának a modellválasztás mellett egyik kulcskérdése kell, hogy legyen.



lődnék.<sup>7</sup> Hasonlóan az előző modellhez, az egyéni és iskolai jellemzők alapján meg lehet becsülni a  $t-1$  és a  $t$  időpont közötti fejlődés elvárható értékét, majd ennek a becsült értéknek és a megfigyelt valós értéknek a különbségét iskolánként átlagolva, kiszámolni az iskola hatását.

### C) „Kombinált” modell

A két modell kombinációjához már nem csupán két, hanem három év tanulói szinten összekötött adataira van szükség:

$$(10) \quad (y_{ijt} - y_{ij(t-1)}) = a_0 + a_1 y_{ij(t-2)} + bX + e_{ij},$$

ahol  $HE_{E3}$  a hibatag iskolai átlaga:

$$(11) \quad HE_{E3} = \hat{e}_{ij}.$$

Ez a modell a fejlődés-modellhez képest azt is feltételezi, hogy a tanulói teljesítménynövekedés üteme nem csupán a családi háttértől (egyéni jellemzőktől) és az iskolai jellemzőktől függ, hanem a tanuló előző teljesítményétől is.

A három modell feltételezéseiben (egyéni szinten normál eloszlású hiba, amely független a bevont kontrollváltozóktól) megegyezik egymással. Értelmezésük is viszonylag egyszerű, adatigényük is hasonló egymáshoz.

### Random hatás-modellek – iskolai random hatás (többszintű modellek)

Az egyszerű lineáris modellek továbbfejlesztett változatai az iskolai random hatást is magukban foglaló modellek. Ezek az úgynevezett többszintű vagy hierarchikus modellek két regressziós egyenletből állnak: egy egyéni és egy iskolai szintű becslésből. Ez annyiban tér el a fentebbi lineáris becsléstől, hogy a konstans tagot iskolák között engedi szóródni, és a hozzáadott érték ( $HE_{it}$ ) az iskolai szintű regresszió (12b) hibatagjának felel meg (13):

$$(12) \quad \begin{aligned} (a) : y_{ijt} &= a_{0j} + a_1 y_{ij(t-1)} + bX + e_{ij}, \\ \text{illetve} \\ (b) : a_{0j} &= A + \delta_{0j}, \end{aligned}$$

<sup>7</sup> Amennyiben az előző modellben az  $a_1 = 1$ , a két modell azonos eredményre vezet.

$$(13) \quad H\hat{E}_{\pi} = \delta_{0j}$$

Az egyenletben becsült két hibateg (az  $e_{ij}$  egyéni és a  $\delta_{0j}$  iskolai) függetlenek egymástól, és mindkettő 0 átlagú fix szórású random tag. Ezeket a modelleket általában „random hatás” modelleknek is hívják, mivel az iskola hatását egy random változó írja le, szemben a fix-hatás modellekkel. A modelleket a fentebb leírt módon lehet „fejlődési” ( $H\hat{E}_{12}$ ) és „kombinált” ( $H\hat{E}_{13}$ ) modellekként is becsülni.

### Fix hatás-modellek – iskolai fix hatás

A random hatással szemben itt az iskola hatását nem egy random, hanem egy fix taggal becsüljük. Vagyis a lineáris regresszióba illesztett iskolai dummy változók koefficiensei lesznek az iskolahatások.

$$(14) \quad y_{ijt} = a_0 + a_1 y_{ij(t-1)} + bX + \Theta_j + e_j,$$

ahol

$$(15) \quad H\hat{E}_{F1} = \Theta_j,$$

vagyis az iskolai fix hatás a hozzáadott érték mutató. Hasonlóan a fentiekhez az iskolai fixhatás-modelleket is lehet „fejlődési” ( $H\hat{E}_{F2}$ ) és „kombinált” ( $H\hat{E}_{F3}$ ) modellekként is becsülni.

### Az egyes modellek összevetése<sup>8</sup>

Gyakorlatban *mindegyik modell használható*, sőt – mint az a tanulmány empirikus fejezetében látható lesz – eredményeiben is hasonlóak egymáshoz, azonban a becslések lényegesen különbözhetnek az egyes iskolára vonatkoztatva. Bár teoretikusan a „kombinált” modellek feltétlenül előnyösebbnek tűnnek, mint a „kiigazított” vagy

<sup>8</sup> A modellek kiválasztásához köthető statisztikai, módszertani problémákat – mint például a kihagyott változók, a hiányzó értékek problémája vagy a modell hatásossága, illetve torzítatlansága – a tanulmány nem tárgyalja. A témáról lásd részletesebben az OECD jelentését (OECD, 2008) vagy a Rand Corporation-nak készített szakértői jelentést (McCaffrey és mtsai., 2003).

a „*fejlődés*”-modellek, kivitelezhetőségük sokkal problematikusabb. Elsősorban három időpontban megfigyelt tanulói szinten összekötött adatbázist sokkal nehezebb (költségesebb és időigényesebb) létrehozni, mint a két időpontban megfigyelt adatokat tartalmazó adatbázisokat. Ám, még ha sikerül is az összekötés, az adatvesztés (évismétlők, lemorzsolódók, válaszmegtagadó stb.) által okozott problémák sokkal nagyobbak, mint a másik két esetben. Ez természetesen további módszertani problémákat is felvet. Ráadásul *Ladd és Walsh* (2002) úgy érvel, hogy ha rendelkezésre áll 3 év tanulói szinten összekötött adatbázisa, akkor a legelső év pontszámait sokkal célszerűbb a második év pontszámainak instrumentumaként felhasználni, és ezáltal szűrni ki a pontszámok – a szerzők eredményei szerint igen jelentős mértékű – mérési hibáját.

A „*kiigazított*” és a „*fejlődés*”-modellek között nincs ilyen egyértelmű átváltás. Ugyanazon az adatbázison meg lehet valósítani mindkettőt, ám a kapott eredmények mást jelentenek. Míg a „*kiigazított*” modellek esetén az adott évben az egyes iskolák tanulóinak átlagos teljesítménye szintbeli különbségét vetjük össze (figyelembe véve, hogy előtte hogyan teljesítettek, illetve milyen egyéni és iskolai jellemzőik vannak), addig a „*fejlődés*”-modellek esetében a két időpont közötti fejlődésük ütemét vetjük össze, és azt feltételezzük, hogy ezt a fejlődési ütemet meghatározzák az egyenletbe bevont egyéni és iskolai változók. Vagyis kontrollváltozók bevonása nélkül a két becslés tökéletesen azonos eredményeket produkálna.

Az iskolai random és fix hatás-modellek nagyon hasonlóak egymáshoz. De míg a fix tag használata esetén a hozzáadott érték mutató pusztán a kontrollként bevont változóktól függ, azaz csak az iskola (és az oda járó diákok) saját jellemzőitől, random hatás használata esetén a modell az összes adatbázisba bevont iskola jellemzőit is figyelembe veszi, amelyet egy úgynevezett „összehúzás” (*shrinking*) segítségével ér el. Ez a legegyszerűbb iskolai random hatás-modellek esetén a következő:

$$(16) \quad \Theta_j = \lambda_j (\hat{y}_j - \bar{y}),$$

ahol  $\hat{y}_j$  az iskola mért tanulói teljesítményének az átlagát jelöli, míg  $\bar{y}$  az országos átlag. A  $\lambda_j$  egy 1-nél kisebb súly, azaz az iskolai átlag és az országos átlag közötti hiba „össze van húzva” nulla felé. A fix hatás-becslés ezzel szemben pontosan a két átlag különbségével egyezik meg (lásd bővebben *McCaffrey és mtsai.*, 2003: p. 65).<sup>9</sup>

A *random*, illetve *fix hatás-modellek* között tehát hasonlóan nehéz a választás (*McCaffrey és mtsai.*, 2003; *OECD*, 2008). Mint az fentebb látható, az iskolai ran-

<sup>9</sup> Míg a random hatás-becslés esetén a  $\hat{y}$  az iskolai átlagok súlyozott átlaga, azt feltételezve, hogy  $\Sigma(\hat{y} - y) = 0$ , addig a fix hatás-becslés esetén  $\hat{y}$  a valós átlag.

dom és fix hatás-modellek között az a különbség, hogy a random hatás szórása, az „összehúzás” miatt, valamelyest kisebb lesz, mint a fix hatásé, ugyanis a random hatás-modellek a nagyobb „hibával” rendelkező iskolákat az átlag felé húzzák. Továbbá minden hozzáadott érték-becslésnél fontos figyelembe venni, hogy a becslés megbízhatósága nagyban függ az iskola méretétől (*Kane és Staiger, 2002*). A kis iskolákra adott becslés mindig megbízhatatlanabb, mint a nagy iskolákéra adott. Így a random hatás-modellekben az iskola mérete miatti eltérések kevésbé jelennek meg. Mivel a random hatás becslése a szélsőségesen jó vagy rossz iskolákat inkább az átlag felé húzza, így várhatóan épp a kisebb iskolák lesznek azok, amelyekre a random hatás-becslés más eredményeket ad. Vagyis míg országos közpolitikai szempontból a random hatás-modellek kisebb hibával becsülnek, az egyes iskolákra nézve a random hatás, az „összehúzás” miatt a becsült érték el fog térni a valós értéktől, azaz torzított lesz a becslés. A random hatás használata mellett szól továbbá, hogy ezeket a modelleket sokkal egyszerűbb úgy kibővíteni, hogy figyelembe vegyenek egyéb iskolai tényezőket is (például random meredekséget, vagy egyéb iskolai szintű jellemzőket), bár ezek fix hatás számításánál is megoldhatók, csak kevésbé egyértelműen.

Összességében a random hatás-becslés akkor lehet jobb, ha a cél közpolitikai jellegű, például ha az iskolák egyes tulajdonságainak az iskola eredményességére való hatását vizsgáljuk. Egyéb esetekben – például amikor az iskolák hatását elszámoltathatósági céllal mérjük – a választás nem egyértelmű, a kisiskolákkal kapcsolatos probléma ugyanis mindkét becslésnél fennáll (lásd például *Kane és Staiger, 2002*). Bár a fix hatás-becslés esetében a becslés és a valós érték közötti eltérés (a hiba) nincs mesterségesen összehúzva, így a becslés nem lesz torzított (várható értékben minden iskola pont olyan becsült értéket kap, amekkora a feltételezett iskolahatás), azonban a kisiskolákban és/vagy a nagyon heterogén iskolákban a becslés így is nagyon bizonytalan lesz, nagy szórással. Vagyis ezekben az esetekben két rossz közül kell választani:

1. vagy a hibát mesterségesen csökkentjük, tudva, hogy az extrém esetekben – a különösen rossz és a különösen jó iskolák esetében – a hozzáadott érték becslése az átlag felé fog torzítani,
2. vagy kockáztatjuk, hogy a becslés bizonytalansága miatt nagyon nagy lesz egy-egy iskola hozzáadott értékében az évek közötti eltérés, függetlenül az iskola valós teljesítményétől.

Az egyéni hibák iskolai átlaga alapján becsült iskolahatás is torzított, ha azt felételezzük, hogy az iskolai fix hatás a torzítatlan közelítés. Mert például, ha a jó tanulók az eredményesebb (azaz a nagyobb fix hatást mutató) iskolákban, míg a rossz

tanulók a kevésbé eredményes iskolákban csoportosulnak, akkor az egyéni hibák átlagán alapuló hozzáadott érték-becslés lefelé torzított lesz a jó, és felfelé torzított lesz a rossz iskolák esetében. Természetesen, ha a képességek szerinti csoportosulás ennek ellenkezője, akkor a torzítás iránya is megváltozik; vagyis a torzítás iránya empirikus kérdés.

A következő fejezet konkrét példákat sorol fel az egyes, itt bemutatott hozzáadott érték-modellek nemzetközi felhasználásáról, illetve olyan, a gyakorlatban is felhasznált modelleket mutat be, amelyek a fentebb felsorolt modelleknek továbbfejlesztett, sokkal adatigényesebb és bonyolultabb verziói. Nyilvánvalóan mindegyik, így az egyszerűbb, gyakorlatban is felhasznált modell is különbözik valamelyest az elméleti modellektől.

### Hozzáadott érték típusú mérésen alapuló modellek a gyakorlatban

Az alábbi fejezet három, gyakorlatban is alkalmazott hozzáadott érték típusú modellt mutat be. A lengyel modell módszertanilag nem nyújt újdonságot, de érdekes látni, hogy Kelet-Európában is kísérleteznek hasonló modellekkel. A karolinai modellek a hozzáadott érték típusú iskolahatás-számítások egyik legkorábban kialakított típusa, amelyet számos tanulmányban elemeztek már. Végül a dallasi modell szintén elég korai, de a karolinai modellekkel szemben nem az egyszerűséget, hanem sokkal inkább a hiteles iskolahatás-becslést tekinti elsődleges célnak.

### Lengyelország

Lengyelországban az elszámoltathatósági rendszer kialakítása még gyerekcipőben jár, de *Jakubowski* (2008) leírása alapján már szépen felvázolható a rendszerhez felhasznált iskolai hozzáadott érték típusú eredményességmérés. A felhasznált modell egy az egyben a „kiigazított” egyéni hibák iskolai átlagán alapuló hozzáadott érték becslésére épül (*lásd (6)-os és (7)-es egyenleteket*). A kimeneti változó a diák által az alsó-középfokú (*gimnazjum*, azaz a 9. osztály vége) képzési szint befejeztével kötelezően megírt kimeneti teszt eredménye, míg a  $t-1$ -edik időpontban mért teljesítmény a 6. osztály után szintén kötelezően megírt teszt. Míg a 9. évfolyam utáni teszt „éles”, azaz eredménye beszámít a középfokú továbbtanulásba, addig a 6. osztály utáni tesztek pusztán informatív jellegűek. Utóbbi eredményei csak az adott diák számára, illetve aggregált formában elérhetőek, és az eredményeknek inkább visszajelzés funkciója van. Bár a közeljövőben a döntéshozók szándékában

áll iskolai és egyéni változók kontrollálása is a modellben, jelenleg sem megfelelő részletezettségű iskolai jellemzők, sem egyéni szintű családi státuszjellemzők nem állnak a kutatók rendelkezésére. A *Jakubowski* (2008) tanulmányában bemutatott modell így csupán a nemet és egy diszlexiát mutató bináris változót tartalmaz egyéni szintű kontrollként.

E modell kiválasztását három érveléssel támasztja alá az erre kijelölt testület (*Jakubowski*, 2008: p. 5). A három általános érv, hogy a modell teoretikusan megállja a helyét: (1) egyszerű megérteni, (2) semleges, vagyis az iskola eredményessége független az oda járó tanulók eredeti pontszámaitól, képességeitől, és (3) végül egyszerű a kivitelezése. *Jakubowski* (2008) tanulmánya mindemellett azzal érvel, hogy a kiválasztott modell eredményeiben nem térne el jelentősen egy nem lineáris, vagy egy fix vagy random hatást felhasználó modelltől, viszont felhívja a figyelmet *Ladd és Walsh* (2002) fentebb említett kritikájára, miszerint a pontszámokat jelentős hibával mérik, így szükséges a korrekciójuk például egy instrumentális becslés segítségével.

### Észak és Dél Karolina (Egyesült Államok)

Az iskolák elszámoltathatóságára Észak- és Dél-Karolinában is a legegyszerűbb, egyéni hibák iskolai átlagán alapuló „*kiigazított*” (Dél-Karolina), illetve „*fejlődés*” (Észak-Karolina) típusú elszámoltathatósági modelleket alkalmaznak (*Clotfelter és Ladd*, 1996; *Ladd és Walsh*, 2002; illetve magyarul *Muraközy és Horn*, 2005).

A) Dél-Karolinában a (6)-os modell egy változatát becslik:

$$(17) \quad y_{ijt} = a_0 + a_1 y_{ij(t-1)}^m + a_2 y_{ij(t-1)}^o + a_3 (y_{ij(t-1)}^m)^2 + a_4 (y_{ij(t-1)}^o)^2 + a_5 (y_{ij(t-1)}^m * y_{ij(t-1)}^o) + e_{ij},$$

$$(18) \quad H\hat{E}_{DK} = \hat{e}_{ij},$$

ahol  $y^m$  a matematika, míg  $y^o$  az olvasás-szövegértés pontszám,  $a$  pedig a becsült koefficiensek.

Az alapmodelltől a dél-karolinai modell két fontos tulajdonságában tér el. Egyrészt közvetlenül nem veszi figyelembe sem az egyéni családi háttér jellemzőit, sem az iskola egyéb tulajdonságait, másrészt nem lineáris hatást feltételez a jelenbeli és a múltbeli teszteredmények között. Az egyéni hatás kontrollálása ellen az szól, hogy ha figyelembe vesszük, akkor explicit módon is kisebb teljesítményt várnak el az alacsonyabb státuszú tanulóktól, mint a magasabb státuszúaktól. Így az egyéni

jellemzők kontrollálása (és különösen az etnikai hovatartozásé) politikailag igen kényes kérdés. A nem lineáris hatás használatát azzal indokolják a modell alkalmazók, hogy a jobb tanulók empirikusan igazolhatóan gyorsabban is tanulnak, azaz a múltbeli iskolai teljesítmény hatása a jelenbeli teljesítményre nem lineáris. Másként fogalmazva, a korábbi képességnek nem feltétlenül ugyanaz a hatása a jelenbeli teljesítményre, vagyis azt teszik fel, hogy a jobban teljesítő tanulók még jobban, míg a rosszabbul teljesítők relatíve még rosszabbul fognak teljesíteni.

Mivel ez a becslés a jobb háttérű tanulókat oktató iskolákat előnyben részesíti, a dél-karolinai döntéshozók az iskolákat öt kvantilisre bontották a tanulók családi háttere alapján. Az iskolákat a kvantiliseken belül hasonlítják össze a hozzáadott értékük alapján. A torzítás logikája a következő: mivel az egyéni teljesítményeket hibával lehet csak becsülni, az a regresszió, amely a  $t$  időpontban becsült teljesítményeket a  $t-1$  időpontban becsült teljesítményekkel közelíti, kétszeresen torzított ugyanabban az irányban, ami a konfidencia intervallumokat a nulla felé torzítja. Minél nagyobb a mérési hiba, annál nagyobb a torzítás, és annál inkább fog a becsült hozzáadott érték a független változóként felhasznált  $t$  időpontban mért teljesítményekkel korrelálni, így a becsült hozzáadott érték annál inkább fog korrelálni az egyéni családi háttérrel is (Clotfelter és Ladd, 1996).

Az iskola hozzáadott értékének a dél-karolinai modellben is  $e_j$ -t feleltetik meg, azzal a különbséggel, hogy nem az egyéni hibák átlagát, hanem annak iskolai mediánját használják. A hivatalos érvelés szerint a medián használata kiküszöböli azt a nem kívánatos ösztönzést, hogy a tanárok esetleg a jó tanulók eredményeinek további javításával emeljék az átlagot a rosszabb tanulók rovására. A kritikusok kiemelik, hogy a medián használata viszont a „széleket vágja le”, azaz arra ösztönzi a tanárokat, hogy a középén lévő tanulókkal foglalkozzon, s így erőforrásokat csoportosítson át a leszakadó vagy a kiemelkedő tanulóktól.

**B) Az észak-karolinai modell a (8)-as alapmodell egy verziója:**

$$(19) \quad (y_{ijt} - y_{ij(t-1)}) = a_0 + a_1 Y1_{ij(t-1)} + bY2_{ij(t-1)} + e_{ij},$$

$$(20) \quad H\hat{E}_{EK} = \hat{e}_{ij}.$$

ahol  $Y1$  a diák előző évi teljesítményének a szintje (*proficiency level*) és  $Y2$  az adott diák tesztpontszámának az állami átlagtól való eltérése.

Az  $Y1$ -re való kontrollálást azzal indokolják a modell megalkotói és a döntéshozók, hogy a magasabb teljesítményszinten lévő tanulók gyorsabban is fejlődnek,

míg az  $Y_2$  az átlaghoz való konvergálást (*regression to the mean*)<sup>10</sup> hivatott korrigálni. Hasonlóan a dél-karolinai modellhez, a családi háterre itt sem kontrollálnak közvetlenül a modellben (Ladd és Walsh, 2002).

### Dallas (Texas, Egyesült Államok)

A dallasi modell talán az egyik legszofisztikáltabb az iskolák elszámoltathatóságára gyakorlatban is alkalmazott modellek közül (Clotfelter és Ladd, 1996; Ladd és Walsh, 2002; OECD, 2008; Webster és Mendro, 1997).

A dallasi rendszer alapja egy kétlépcsős, „kiigazított” random iskolahatás-becslés, ahol az első lépcsőben (21) a nyers pontszámokat számos egyéni és iskolai/körzet szintű jellemzőtől „megtisztítják”:

$$(21) \quad y_{ijt} = a_0 + b_1 X_{1ijt} + \dots + b_p X_{pijt} + e_j.$$

Az  $X$  itt a gyerek nemét, etnikai hovatartozását, anyagi helyzetét, ezek első és második tagú interakcióit – azaz a változók egymásra gyakorolt kereszthatásait (szorzatait) – és néhány iskolakörzeti változót jelent.

A második lépcsőben az első lépcső becsült hibáját standardizált formában használják fel (legyen  $\sim$  a standardizálás jele). A második lépcső egy többszintű, random hatás-becslés, ahol nem csupán a diákok múltbeli teljesítményére kontrollálnak (2 évre visszamenően), hanem további iskolai szintű jellemzőket is bevonnak az egyenletbe:

$$(a): \tilde{e}_{ijt} = c_{0j} + c_{1j} y_{ij(t-1)} + c_{2j} y_{ij(t-2)} + \delta_{ij}$$

*illetve*

$$(22) \quad (b): c_{0j} = G_{00} + \sum_{k=1}^m G_{0k} W_{0k} + u_{0j}$$

$$(c): c_{1j} = G_{10} + \sum_{k=1}^m G_{1k} W_{1k}$$

$$(d): c_{2j} = G_{20} + \sum_{k=1}^m G_{2k} W_{2k},$$

<sup>10</sup> Átlaghoz való konvergálásnak azt az empirikus megfigyelést hívjuk, hogy az első időszakban jól teljesítők a második időszakban általában relatíve rosszabbul, míg a rosszul teljesítők relatíve jobban teljesítenek. Ennek számos oka lehet, a mérési hibától kezdve, az egyéni pszichikai („múltkor jól ment, most nem figyelek annyira”) okokig bezárólag.



ahol  $c$ , illetve  $G$  becsült koefficienseknek felelnek meg, míg  $W$  *m darab* olyan iskolai jellemző, ami az iskola intézkedéseitől független (*exogén*), mint például az iskola néhány demográfiai jellemzője, az iskolába járó diákok társadalmi státuszára vonatkozó változók vagy az iskolába járó diákok fluktuációja vagy az iskolai túlszűfolttság mutatója. A számított iskolahatás így az  $u_{0j}$  lesz:

$$(23) \quad H\hat{E}_{Dallas} = u_{0j}.$$

A dallasi modell felfogásában szinte gyökeres ellentéte a tervezett lengyel modellnek. Eredményeit nehéz még egy képzett szakember számára is megérteni, bonyolult és adatigényes a kivitelezhetősége, viszont feltehetőleg jelenleg a politikailag legkorrektebb modell, amely a legkevésbé enged teret a módszertani hibáknak, kontrollál a legtöbb ismert iskolák döntésein kívül eső tényezőre és még alkalmas a gyakorlati politikai következtetések levonására is.

### Országos kompetenciamérés<sup>11</sup>

Jelenleg Magyarországon minden iskola, minden telephely és iskolafenntartó évente kap egy rövid összefoglaló értékelést a tanulók összesített teljesítményéről. Ezek a jelentések mindig az Országos kompetenciamérés (OKM) előző évi adatain alapulnak. A kompetenciamérés több oktatáspolitikai célt is szolgál. Egyrészt a központi kormányzatot hivatott tájékoztatni az országos, regionális és helyi oktatás színvonaláról, illetve egyéb jellemzőiről. Másrészt az iskolák, telephelyek vezetőit és az iskolafenntartókat tájékoztatja az egyes iskolák teljesítményéről, az előző év adatai alapján. Harmadrészt adalékul szolgál a tanároknak a mérési-értékelési gyakorlatuk fejlesztéséhez, végül a szülők is tájékozódhatnak a gyermekük és az iskolák eredményéről.

Az Országos kompetenciamérést 2001-ben kezdte el megvalósítani a központi oktatási kormányzat, azóta összesen 6 alkalommal mérték fel teljes körűen a 6., 8. és 10. évfolyamos tanulók olvasási és matematikai műveltségét (lásd 2. táblázat).

<sup>11</sup> Az OKM felméréséről lásd bővebben Hermann és Molnár (2008), valamint Balácsi és munkatársai (2014) tanulmányát, illetve e kötetben Ostorics László írását.

## 2. táblázat

Az Országos kompetenciamérés adatbázisainak néhány jellemzője

Év	6. évfolyam	8. évfolyam	10. évfolyam
2003	iskolánként 20 fő	–	iskolánként és képzési típusonként 20 fő
2004	iskolánként 20 fő	iskolánként 20 fő	iskolánként és képzési típusonként 20 fő
2006	195 országosan reprezentatív intézményben teljes körű	teljes körű	minden iskola minden telephelyéről képzési formánként 30 fő
2007	200 országosan reprezentatív intézményben teljes körű	teljes körű	minden iskola minden telephelyéről képzési formánként 30 fő
2008*	teljes körű	teljes körű	teljes körű
2009*	teljes körű	teljes körű	teljes körű
2010*	teljes körű	teljes körű	teljes körű
2011*	teljes körű	teljes körű	teljes körű
2012*	teljes körű	teljes körű	teljes körű
2013*	teljes körű	teljes körű	teljes körű
2014*	teljes körű	teljes körű	teljes körű

\*Oktatási azonosítón alapuló egyéni mérési azonosítók alapján felvett mérések.

A kompetenciamérés, azon kívül, hogy egy kidolgozott, több órás kérdőívvel felméri a tanulók olvasási és matematikai műveltségét, tartalmaz egy tanulói háttérkérdőívet, illetve egy telephelyi és egy iskolai szintű kérdőívet is. A tanulói kérdőív a tanulók, illetve a szüleik legfontosabb szociokulturális és gazdasági jellemzőit méri fel, a telephelyi és az iskolai kérdőívből készült adatbázisok pedig az adott telephelyek és iskolák legfontosabb erőforrás-ellátottságáról tartalmaznak információkat.

Az iskolajelentések ezen adatokon alapulnak. Az úgynevezett FIT (Fenntartói, Intézményi, Telephelyi) jelentések 2006 óta iskolánként szabadon hozzáférhetőek minden érdeklődő számára.<sup>12</sup> A jelentések alapvetően 3 típusú információt tartalmaznak:

1. megmutatják az iskolák átlageredményét mindkét mért műveltségterületen (olvasás-szövegértés és matematika),
2. megmutatják az iskolák tanulóinak képességeloszlását (azaz, hogy a tanulók hány százaléka van az egyes képességszinteken), illetve

<sup>12</sup> <https://www.kir.hu/okmfit>

3. ezeket az adatokat összevetik az adott településtípus és a képzés jellege szerinti országos átlaggal.

Az iskolai jelentések ezeken az információkon túl megmutatják az iskola elvárható eredményét is. Az elvárható eredményt a következő egyenletből becslik:

$$(24) \quad y_{ij} = a_0 + bX + e_{ij},$$

ahol  $i$  a diákot,  $j$  az iskolát,  $y$  a diák mért teljesítményét jelöli, míg  $X$  az egyéni és iskolai jellemzők (a kontrollváltozók) vektora. Az  $X$  vektor tartalmazza a tanulók családi háttérindexét (CSHI), illetve közvetve az iskola típusát és településének típusát, amennyiben ezeket összeveti az adott képzési-, illetve településtípusú intézmények eredményeivel.<sup>13</sup> Az  $a_0$  és  $a_1$ , illetve  $b$  vektor becsült koefficiensek, míg  $e$  az egyéni hiba jele.

Az elvárható eredményt ( $\hat{y}_{ij}$ ) az iskolajelentés összeveti a tényleges eredménnyel ( $y_{ij}$ ), és bár a jelentés közvetlenül nem számolja ki az elvárt eredménytől való eltérést, vagy iskolahatást, a tanulmány további részében ezt vetem össze a hozzáadott érték típusú mutatókkal:<sup>14</sup>

$$(25) \quad IH = \text{átlag}_j (y_{ij} - \hat{y}_{ij}) = e_j.$$

### Hozzáadott érték típusú mutatók az OKM-adatokon

A tanulmányban felsorolt elméleti modelleket a magyar Országos kompetenciamérés adatain is ki lehet számolni. Az alábbiakban a legegyszerűbb alapmodelleket vetem össze egymással (lásd 3. táblázat).

<sup>13</sup> Csak abban az esetben számol a jelentés elvárt eredményt, ha a tanulók legalább kétharmadának van érvényes családi háttérindexe. A telephelyi jelentés az átlageredmények, a képességeloszlás és az elvárt eredmények mellett az átlageredmények változását is megmutatja az előző évi eredményekhez viszonyítva.

<sup>14</sup> A tanulmányban használt iskolahatás a tanulók tényleges és elvárt eredményének iskolai átlaga (vagyis az egyéni hibát felhasználó modell), míg a jelentésekben a tanulók tényleges eredményének iskolai átlaga és az elvárt eredményének iskolai átlaga van megadva; vagyis ez utóbbi nem egyezik meg pontosan a tanulmány által használt, keresztmetszeti adatokon becsült iskolahatással. Ennek ellenére úgy gondolom, hogy a tanulmány által használt mutató jobban összevethető a többi, az irodalomban megjelenő hozzáadott érték típusú mutatóval.

**3. táblázat**  
A különböző modellek rövidítései

Rövidítés	Iskolahatás
<b>Keresztmetszeti adatokon becsült mutatók</b>	
<b>IHe*</b>	Keresztmetszeti adatokon becsült iskolahatás (tényleges – CSHI alapján elvárt eredmény), egyéni hiba
<b>IHi</b>	Keresztmetszeti adatokon becsült iskolahatás (tényleges – CSHI alapján elvárt eredmény), iskolai random hatás
<b>IHf</b>	Keresztmetszeti adatokon becsült iskolahatás (tényleges – CSHI alapján elvárt eredmény), iskolai fix hatás
<b>Hozzáadott érték típusú mutatók</b>	
<b>HEe1</b>	Egyéni hiba, „kiigazított”
<b>HEi1</b>	Iskolai random hatás, „kiigazított”
<b>HEf1</b>	Fix hatás, „kiigazított”
<b>HEe2</b>	Egyéni hiba, „fejlődés”
<b>HEi2</b>	Iskolai random hatás, „fejlődés”
<b>HEf2</b>	Fix hatás, „fejlődés”
<b>HEe3</b>	Egyéni hiba, „kombinált”
<b>HEi3</b>	Iskolai random hatás, „kombinált”
<b>HEf3</b>	Fix hatás, „kombinált”

\*A FIT-jelentésekben számolt elvárható eredményhez leginkább hasonló mutató.

Jelen számítás csupán egy személtetése annak, hogy mennyiben befolyásolja az egyes iskola értékelését a módszertan. Fontos hangsúlyozni, hogy az elméleti modellek egyike sem jobb vagy rosszabb, mint a másik. Különbség köztük csupán az alapfeltevésekben van. A modellek az előző időszak(ok)ban mért tesztpontszámok mellett csupán a családi háttérindexre kontrollálnak<sup>15</sup>: nem veszik figyelembe a tanulók egyéb jellemzőit, az iskola típusát vagy egyéb iskolai jellemzőket, vagyis semmiképp nem tekinthetők végleges eredményességmutatóknak.

A 2012-es évre matematikából és szövegértésből számolt páronkénti korrelációs együtthatókat vizsgálva jól látható, hogy a keresztmetszeti adatokon alapuló elvárt eredmény és valós eredmény különbsége (IHe, IHi, IHf) leginkább a „kiigazított” becslésekkel (IHe1, IHi1, IHf1) korrelál. A korrelációs együttható minden esetben 0,7 felett van, de egyes mutatók között nagyon szoros (0,96) együtt járás is tapasztalható. Vagyis a FIT-jelentésekben megjelenő iskolahatás-mutatók (a családi háttérindexre kontrolláló modellek) nem térnek el alapjaiban a FIT-jelentések „komplex modelljeitől”. A FIT-jelentések komplex modelljei megfeleltethetők az itt bemutatott

<sup>15</sup> A keresztmetszeti adatokon becsült modellek (IHe, IHi, IHf) csak a családi háttérindexre kontrollálnak.

„kiigazított”, egyéni hibát használó modelleknek. Ez nem jelenti azt, hogy egy-egy iskolára nézve ne adna más eredményt az egyéni hibát használó iskolahatás (IHe) vagy a „kiigazított” egyéni hibát használó (HEe1) modell (lásd 4. táblázat).

4. táblázat

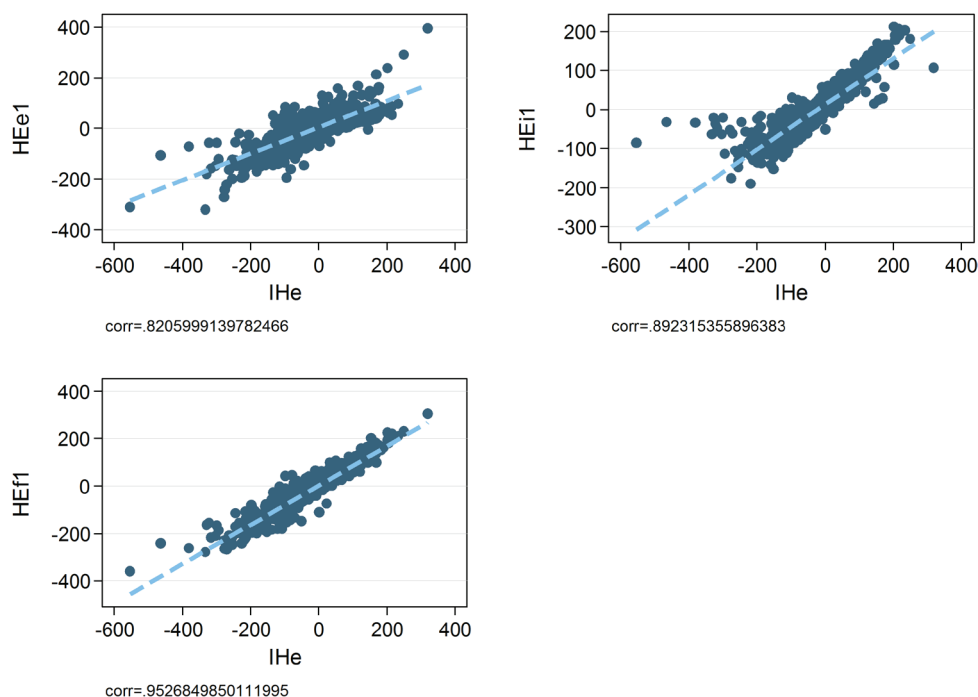
Páronkénti korrelációs együtthatók az iskolahatások egyes mutatói között (2012)

Matematika	IHe	IHi	IHf	HEe1	HEi1	HEf1	HEe2	HEi2	HEf2	HEe3	HEi3	HEf3
IHe	1,00	–	–	–	–	–	–	–	–	–	–	–
IHi	0,91	1,00	–	–	–	–	–	–	–	–	–	–
IHf	0,95	0,97	1,00	–	–	–	–	–	–	–	–	–
HEe1	0,82	0,70	0,75	1,00	–	–	–	–	–	–	–	–
HEi1	0,89	0,96	0,92	0,76	1,00	–	–	–	–	–	–	–
HEf1	0,95	0,93	0,96	0,84	0,95	1,00	–	–	–	–	–	–
HEe2	0,48	0,37	0,41	0,89	0,47	0,54	1,00	–	–	–	–	–
HEi2	0,49	0,45	0,45	0,84	0,56	0,56	0,91	1,00	–	–	–	–
HEf2	0,50	0,42	0,46	0,90	0,50	0,57	1,00	0,92	1,00	–	–	–
HEe3	0,50	0,39	0,44	0,90	0,48	0,56	0,98	0,90	0,98	1,00	–	–
HEi3	0,52	0,48	0,48	0,85	0,58	0,59	0,90	0,99	0,91	0,91	1,00	–
HEf3	0,54	0,45	0,50	0,91	0,53	0,60	0,98	0,91	0,98	1,00	0,92	1,00
Szövegértés	IHe	IHi	IHf	HEe1	HEi1	HEf1	HEe2	HEi2	HEf2	HEe3	HEi3	HEf3
IHe	1,00	–	–	–	–	–	–	–	–	–	–	–
IHi	0,91	1,00	–	–	–	–	–	–	–	–	–	–
IHf	0,95	0,96	1,00	–	–	–	–	–	–	–	–	–
HEe1	0,83	0,71	0,77	1,00	–	–	–	–	–	–	–	–
HEi1	0,87	0,96	0,92	0,74	1,00	–	–	–	–	–	–	–
HEf1	0,92	0,91	0,96	0,82	0,94	1,00	–	–	–	–	–	–
HEe2	0,47	0,36	0,41	0,88	0,43	0,51	1,00	–	–	–	–	–
HEi2	0,44	0,41	0,40	0,78	0,51	0,48	0,86	1,00	–	–	–	–
HEf2	0,49	0,40	0,45	0,89	0,47	0,55	1,00	0,86	1,00	–	–	–
HEe3	0,47	0,37	0,42	0,87	0,44	0,53	0,98	0,84	0,98	1,00	–	–
HEi3	0,37	0,32	0,32	0,74	0,43	0,41	0,86	0,98	0,86	0,86	1,00	–
HEf3	0,42	0,32	0,37	0,84	0,40	0,49	0,97	0,82	0,97	0,99	0,85	1,00

Az 1. ábra bal felső ábráján jól látható, hogy egyes iskolák jóval az azonosságot mutató szaggatott vonal alatt vagy felett helyezkednek el. Minél távolabb esik egy iskola a szaggatott vonaltól, annál jobban eltér a két mutatóból levonható következtetés az adott iskola esetén.

1. ábra

A „kiigazított” egyéni és iskolai random hatás, illetve a fix hatás-modellek összevetése az egyéni hibát felhasználó, keresztmetszeti adatokon becsült iskolahatás-mutatókkal (OKM 2012/10. évf.–2010/8. évf., matematika)

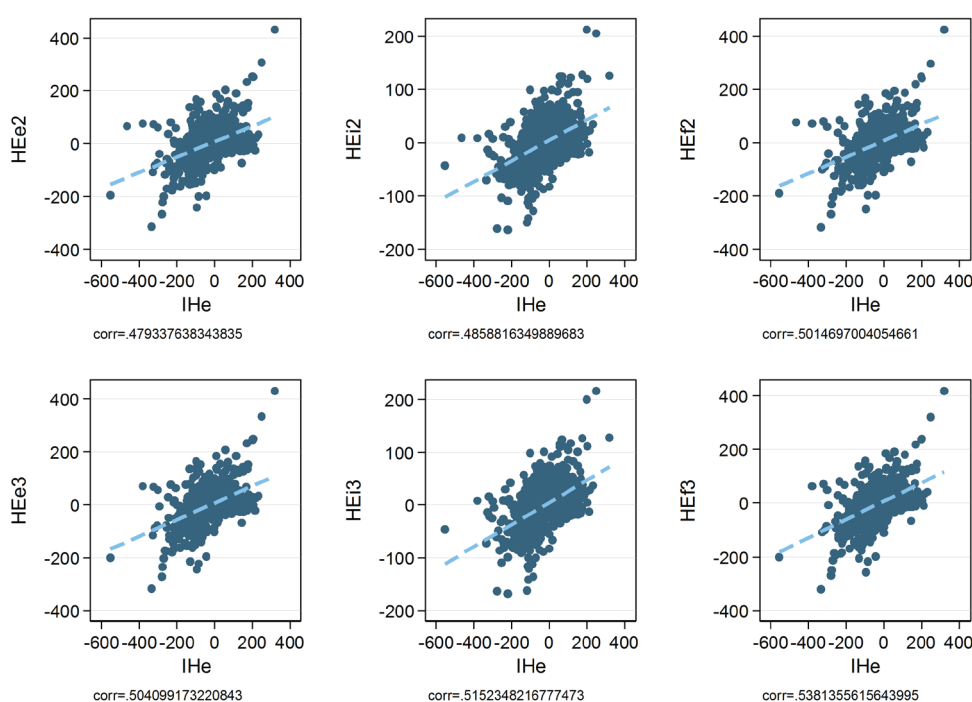


A 4. táblázatból az is leolvasható, hogy a FIT-jelentések által közölt elvárható értékhez leginkább hasonló egyéni hibát használó, keresztmetszeti adatokon becsült iskolahatás-mutatók együtt járása már korántsem olyan szoros, ha a „fejlődés” vagy a „kombinált” hozzáadott érték típusú modellekkel vetjük őket össze. A korrelációs együtthatók itt szinte sehol nem haladják meg a 0,5-ös értéket. Ráadásul, ha nem is az egyéni hibát használjuk a keresztmetszeti adatokon becsült iskolahatás kiszámításához, hanem iskolai random hatást (IH<sub>i</sub>), és ezt vetjük össze a szintén iskolai random hatás modellel kiszámolt „kombinált” modellel (HE<sub>i3</sub>), a korrelációs együttható 0,32-re csökken a szövegértés esetén. Tekintettel arra, hogy mindkét mutató azzal a céllal jött létre, hogy az iskolák eredményességéről adjon tájékozta-

tást, ez az együtt járás kifejezetten gyengének mondható. Jól látható, hogy számos iskola jóval a szaggatott vonal alatt vagy felett van (lásd 2. ábra). Ezen iskolákra nézve különösen nagy tétje lenne annak, ha az egyik vagy a másik számítási módszer szerint lennének értékelve.

2. ábra

Az „fejlődés”, illetve a „kombinált” egyéni és iskolai random hatás, illetve a fix hatás-modellek összevetése az egyéni hibát felhasználó, keresztmetszeti adatokon becsült iskolahatás-mutatókkal (OKM 2012/10. évf.–2010/8. évf., matematika)

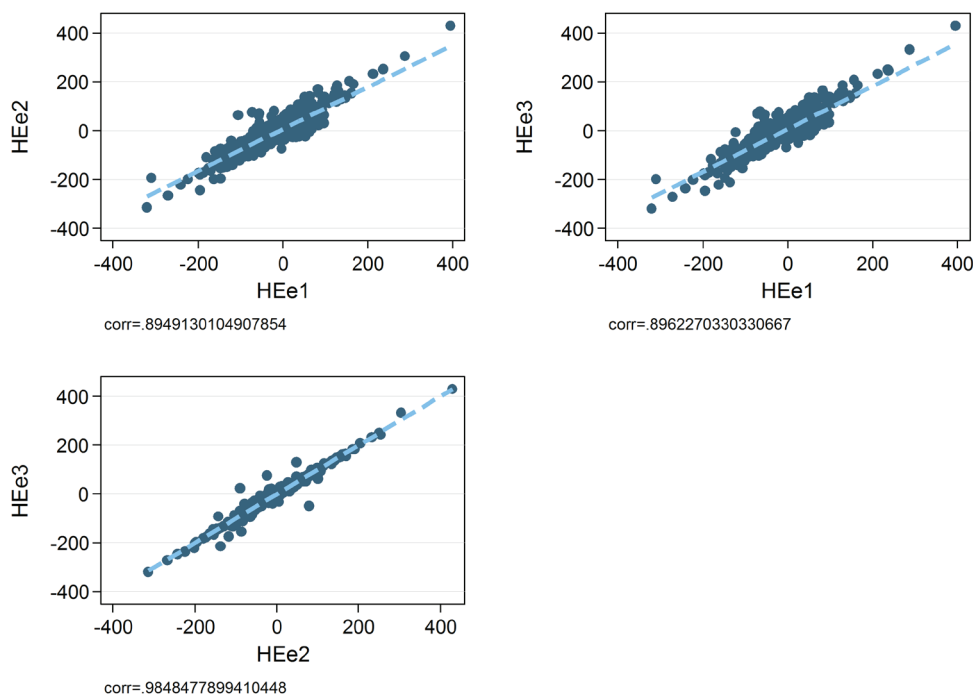


A 3. ábrán az egyes hozzáadott érték típusú – vagyis a korábbi tesztpontszámokra és a családi háttérindexre is kontrolláló – mutatók, a „kiigazított”, a „fejlődés” és a „kombinált” modellek egyéni hibát felhasználó eredményei vannak összevetve egymással. Jól látható, hogy önmagában az, hogy milyen modellt feltételezünk, nem határozza meg jelentősen az egyes iskolák helyzetét: a korrelációs együtthatók mindenütt 0,9 körüliek, vagy e feletti. Vagyis úgy tűnik, a másik dimenzió sokkal inkább meghatározó: az, hogy egyéni hiba, iskolai random hatás vagy iskolai fix hatás-modellekkel számolják az iskolák eredményességét, nagyobb mértékben meghatározza az egyes iskolák helyzetét, mint az, hogy a múltbeli

teszteredmény(ek)e)t független változóként („*kiigazított modell*”), függő változóként („*fejlődés modell*”) vagy mindkét helyen („*kombinált modell*”) figyelembe veszik.

### 3. ábra

A „kiigazított”, „fejlődés” és „komplex” hozzáadott érték modellek egyéni hibát felhasználó eredményein alapuló mutatók összevetése (OKM 2012/10. évf.–2010/8. évf., matematika)



Az eredmények azonban még egy dimenzióban óvatosságra intenek: az egyes módszerek által számított eredményességmutatók jelentősen eltérhetnek egymástól az évek között. Jól látható, hogy a 2012. és a 2011. évre ugyanolyan módon számolt mutatók korrelációs együtthatói 0,5–0,8 között mozognak (lásd 5. táblázat).



## 5. táblázat

Páronkénti korrelációs együtthatók a „kiigazított” hozzáadott érték modellek között  
2011-ben és 2012-ben\*

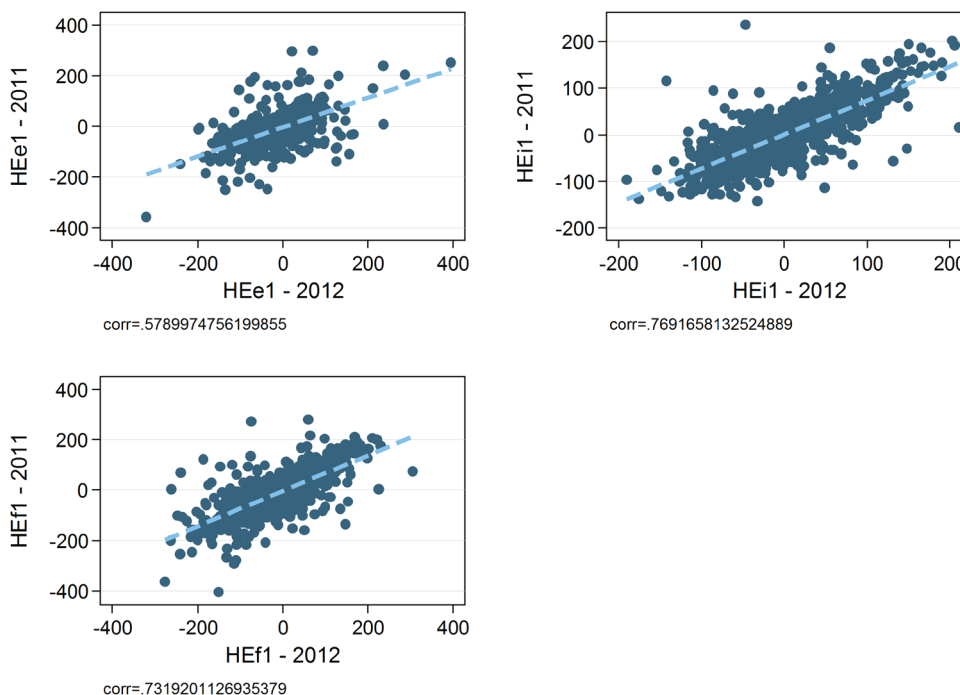
Matematika		2011			2012		
		HEe1	HEr1	HEf1	HEe1	HEr1	HEf1
2011	HEe1	1,00	–	–	–	–	–
	HEr1	0,79	1,00	–	–	–	–
	HEf1	0,87	0,95	1,00	–	–	–
2012	HEe1	<b>0,58</b>	0,54	0,57	1,00	–	–
	HEr1	0,53	<b>0,77</b>	0,71	0,76	1,00	–
	HEf1	0,58	0,74	<b>0,73</b>	0,84	0,95	1,00
Szövegértés		2011			2012		
		HEe1	HEr1	HEf1	HEe1	HEr1	HEf1
2011	HEe1	1,00	–	–	–	–	–
	HEr1	0,73	1,00	–	–	–	–
	HEf1	0,80	0,94	1,00	–	–	–
2012	HEe1	<b>0,46</b>	0,52	0,52	1,00	–	–
	HEr1	0,50	<b>0,82</b>	0,76	0,74	1,00	–
	HEf1	0,49	0,76	<b>0,75</b>	0,82	0,94	1,00

\*A félkövérrel jelölt számok a különböző évek adatain, azonos módon számított mutatókat jelölik.

Mivel nem valószínű, hogy az egyes iskolák *valódi* eredményessége két egymást követő évben nagymértékben változna, így feltételezhető, hogy a hozzáadott érték típusú mutatók sem teljesen függetlenek olyan egyéb hatásoktól, mint amilyen például a kohorsz-hatás, hiszen az egyetlen fontos eltérés a két időpontra számolt mutatók között, hogy más tanulók szerepelnek benne. Az egyikben azok, akik 2012-ben voltak 10. évfolyamosok, míg a másokban azok, akik egy évvel korábban, 2011-ben. Ennek megfelelően az egyes évek között feltehetően éppen a kohorsz-hatás miatt jelentősen változik/változhat egy-egy iskola hozzáadott érték mutatóval mért eredményessége (lásd 4. ábra).

4. ábra

A különböző „kiigazított” hozzáadott érték típusú mutatók összevetése 2011 és 2012 között



Ezekből a megfigyelésekből azt a következtetést lehet levonni, hogy érdemes az iskolák eredményességét nem egy év, hanem több év összevont eredményessége alapján vizsgálni. Hogy ez az összevonás miként történik, feltehetően szintén jelentősen befolyásol(hat)ja az egyes iskolák hosszabb távú megítélését.

### Összefoglalás

A tanulmány az elméletileg lehetséges hozzáadott típusú indikátorok alapverzióit mutatja be, és nemzetközi példákat is hoz ezekre. A tanulmány kísérletet tesz arra is, hogy magyar adatokon megvizsgálja, valóban más eredményeket kapnának-e az iskolák, ha nem keresztmetszeti, hanem több időpontban mért, tanulói szinten összekötött adatokon becsült, hozzáadott érték típusú eredményességi mutatókat kapnának kézhez. Fontos hangsúlyozni, hogy a kutatás egy *pilot*, amely nem alkalmas iskolai szintű következtetések levonására, de az oktatáspolitikai döntéshozók számára kulcsfontosságú információkkal szolgálhat.

A tanulmány két fontos dologra hívja fel a figyelmet: egyrészt arra, hogy a keresztmetszeti adatokon nyugvó statisztikák nagy valószínűséggel torzított képet fognak nyújtani egy adott iskola valós eredményességéről. E torzítás mértékének a csökkentésére lehetőség van hozzáadott érték típusú indikátorok figyelembevételével. A tanulmány másik fontos állítása azonban az, hogy a hozzáadott érték típusú mutatók sem tökéletesek. Egyrészt nem oldják meg, csak csökkentik a mérésből fakadó bizonytalanságot, másrészt nem létezik egy tökéletes „hozzáadott érték mutató”, hanem sokféle, különböző feltevéseken alapuló *hozzáadott érték típusú mutató* létezik, amelyek adott esetben más és más eredményeket mutatnak egy adott iskolára nézve.

A tanulmány fő konklúziója, hogy gyakorlatban mindegyik modell használható, sőt eredményeiben is hasonlóak egymáshoz, azonban a becslések lényegesen különbözhetnek az egyes iskolára vonatkoztatva. Még a legegyszerűbb hozzáadott érték típusú modellek is jelentősen megváltoztatnák az egyes iskolák eredményességi mutatóit, ha azonban kicsit bonyolultabb, ám az iskolai munka értékelése szempontjából korrektebb mutatókat alkalmaznánk, az eredmények több esetben drasztikusan eltérnének. Mindezek mellett – bár a hozzáadott érték típusú mutatók közti azonos éven belüli korreláció sem tökéletes –, ha két különböző évet hasonlítunk össze, akkor az eltérés már jelentős. Így érdemes egyrészt többféle módszerrel, másrészt több évre kiszámolni az iskolák eredményességét, s így talán egy lépéssel közelebb kerülünk az iskolák valós eredményességének becsléséhez.

### Hivatkozások

- Balázsi Ildikó és mtsai. (2014): *Az Országos kompetenciamérés tartalmi keretei – szövegértés, matematika, háttérkérdőívek*. Oktatási Hivatal, Budapest.
- Clotfelter, C. T. és H. F. Ladd (1996): Recognizing and Rewarding Success in Public Schools. In: H. F. Ladd (Ed.): *Holding Schools Accountable: Performance-based Reform in Education* (pp. 23–63), The Brookings Institution, Washington D. C.
- Dolton, P. J. (2002). Improving Educational Quality: How Best to Evaluate Our Schools? (Discussion). In Y. K. Kodryzcki (Ed.): *Education in the 21st Century* (pp. 225–235), Federal Reserve Bank of Boston.
- Hermann Zoltan és Molnár Tímea Laura (2008): *Országos kompetenciamérési adatbázis*. Elérhető: <http://adatbank.mtaki.hu/files/dokum/7.pdf> [Letöltés dátuma: 2015.04.30.]
- Horn Dániel (2011): Az oktatási elszámoltathatósági rendszerek elmélete. *MTA-KTI Műhelytanulmányok*, no. 13.
- Jakubowski, M. (2008): Implementing Value-Added Models of Schools Assessment. *EUI Working Papers RSCAS*, no. 6.
- Kane, T. J. és Douglas, O. S. (2002): The Promise and Pitfalls of Using Imprecise School Accountability Measures. *The Journal of Economic Perspectives* 16 (November), 91–114.

- Ladd, H. F. és Randall, P. W. (2002): Implementing Value-Added Measures of School Effectiveness: Getting the Incentives Right. *Economics of Education Review*, 21(1), 1–17.
- McCaffrey, D. F. és mtsai. (2003): *Evaluating Value-Added Models for Teacher Accountability. Monograph*. RAND Corporation.
- Muraközy Balázs és Horn Dániel (2005): Teljesítményalapú ösztönzőrendszerek a közoktatásban. In: Hermann Zoltán (szerk.): *Hatékonysági problémák a közoktatásban* (pp.165–185), Országos Közoktatási Intézet, Budapest.
- OECD (2008): *Measuring Improvements in Learning Outcomes – Best Practices to Assess the Value-Added of Schools*. OECD Publishing, Paris.
- Todd, P. E. és K. I., Wolpin (2003): On the Specification and Estimation of the Production Function for Cognitive Achievement. *The Economic Journal*, 113(485), 3–33.
- Webster, W. J. és R. R., Mendro (1997): *The Dallas Value-Added Accountability System*. Dallas Public Schools.